Title: The efficacy of single-trial multisensory memories.

Authors: Thelen A, Murray MM

Journal: Multisensory research

Year: 2013

Issue: 26

Volume: 5

Pages: 483-502

Multisensory Research

The efficacy of single-trial multisensory memories

--Manuscript Draft--

Manuscript Number:	MSR-1021R3
Full Title:	The efficacy of single-trial multisensory memories
Article Type:	Special Issue: 13th International Multisensory Research Forum
Keywords:	multisensory; cross-modal; auditory; visual; object; memory; learning; encoding; retrieval
Corresponding Author:	Micah M. Murray University Hospital Center and University of Lausanne Lausanne, SWITZERLAND
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University Hospital Center and University of Lausanne
Corresponding Author's Secondary Institution:	
First Author:	Antonnia Thelen
First Author Secondary Information:	
Order of Authors:	Antonnia Thelen
	Micah M. Murray
Order of Authors Secondary Information:	
Abstract:	This review article summarizes evidence that multisensory experiences at one point in time have long-lasting effects on subsequent unisensory visual and auditory object recognition. The efficacy of single-trial exposure to task-irrelevant multisensory events is its ability to modulate memory performance and brain activity to unisensory components of these events presented later in time. Object recognition (either visual or auditory) is enhanced if the initial multisensory experience had been semantically congruent and can be impaired if this multisensory pairing was either semantically incongruent or entailed meaningless information in the task-irrelevant modality, when compared to objects encountered exclusively in a unisensory context. Processes active during encoding cannot straightforwardly explain these effects; performance on all initial presentations was indistinguishable despite leading to opposing effects with stimulus repetitions. Brain responses to unisensory stimulus repetitions differ during early processing stages (~100ms post-stimulus onset) according to whether or not they had been initially paired in a multisensory context. Plus, the network exhibiting differential responses varies according to whether or not memory performance is enhanced or impaired. The collective findings we review indicate that multisensory associations formed via single-trial learning exert influences on later unisensory processing to promote distinct object representations that manifest as differentiable brain networks whose activity is correlated with memory performance. These influences occur incidentally, despite many intervening stimuli, and are distinguishable from the encoding/learning processes during the formation of the multisensory associations. The consequences of multisensory interactions thus persist over time to impact memory retrieval and object discrimination.
Suggested Reviewers:	Charles Spence Charles.Spence@psy.ox.ac.uk
	Mark Wallace mark.wallace@vanderbilt.edu
	Durk Talsma durktals@gmail.com
	Kirsten Taylor

	kirtsten.taylor@usb.ch
Opposed Reviewers:	

The efficacy of single-trial multisensory memories

Antonia Thelen[1], Micah M. Murray[1-4]

[1] The Functional Electrical Neuroimaging Laboratory, Neuropsychology and Neurorehabilitation Service, Department of Clinical Neurosciences, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Switzerland
[2]Electroencephalography Brain Mapping Core, Center for Biomedical Imaging (CIBM), Lausanne, Switzerland
[3]Radiology Department, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Switzerland
[4]Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

Address correspondence to:
Prof. Micah M. Murray
The Neuropsychology and Neurorehabilitation Service
Department of Clinical Neurosciences
Centre Hospitalier Universitaire Vaudois
BH08.078
Rue du Bugnon 46
1011 Lausanne, Switzerland
Micah.murray@chuv.ch

Abstract

This review article summarizes evidence that multisensory experiences at one point in time have long-lasting effects on subsequent unisensory visual and auditory object recognition. The efficacy of single-trial exposure to task-irrelevant multisensory events is its ability to modulate memory performance and brain activity to unisensory components of these events presented later in time. Object recognition (either visual or auditory) is enhanced if the initial multisensory experience had been semantically congruent and can be impaired if this multisensory pairing was either semantically incongruent or entailed meaningless information in the task-irrelevant modality, when compared to objects encountered exclusively in a unisensory context. Processes active during encoding cannot straightforwardly explain these effects; performance on all initial presentations was indistinguishable despite leading to opposing effects with stimulus repetitions. Brain responses to unisensory stimulus repetitions differ during early processing stages (~100ms post-stimulus onset) according to whether or not they had been initially paired in a multisensory context. Plus, the network exhibiting differential responses varies according to whether or not memory performance is enhanced or impaired. The collective findings we review indicate that multisensory associations formed via single-trial learning exert influences on later unisensory processing to promote distinct object representations that manifest as differentiable brain networks whose activity is correlated with memory performance. These influences occur incidentally, despite many intervening stimuli, and are distinguishable from the encoding/learning processes during the formation of the multisensory associations. The consequences of multisensory interactions thus persist over time to impact memory retrieval and object discrimination.

Keywords: Multisensory, cross-modal, auditory, visual, object, memory, learning

1. Background

Studies of multisensory object processing have generally concentrated on how information from one sensory modality impacts behaviour and/or brain responses to information from another sensory modality (e.g. Amedi et al., 2005; Naci et al., 2012; Cappe et al., 2010, 2012; Murray et al. 2012; Beer et al., 2013). Substantially less consideration has been given to how multisensory information processing at one point in time affects subsequent (unisensory) processing. This kind of situation is commonplace. After meeting someone, say in the context of a cocktail party, you later recognize her face at a conference or his voice when calling on the telephone. Likewise, acquiring fluent reading skills initially involves ascribing sounds to written letters, but later progresses to whole-word reading without the need for (sub)vocalization. In this review, we provide a summary of recent efforts to describe the perceptual consequences of and spatio-temporal brain dynamics mediating the influence of past multisensory experiences on current unisensory object discrimination.

The few human neuroimaging investigations that have examined how experiences in one or multiple senses alter later processing of stimuli of another sensory modality provide evidence that brain regions involved in an experience's encoding can also be involved during its subsequent active retrieval (e.g. James et al., 2002; Nyberg et al., 2000; Wheeler et al., 2000; see also von Kriegstein and Giraud, 2006). In these studies, subjects learned auditory-visual or visual-visual associations during a separate session and later classified visual stimuli according to the sensory modality with which it was initially paired (Nyberg et al., 2000; Wheeler et al., 2000). During a test session auditory regions were active in response to those visual stimuli that had been presented with sounds during study sessions. This activity was taken as support for the psychological postulate of 'redintegration' (Hamilton, 1859), wherein a component part is sufficient to (re)activate the whole experience's consolidated representation. That is, a visual stimulus that had been studied and thus associated with a sound (and presumably formed a consolidated representation with that sound) could elicit activity within auditory cortices when participants actively remembered the initial encoding context.

Intracranial microelectrode recordings in monkeys during the performance of a delayed match-to-sample task provide a similar line of evidence (e.g. Colombo and Gross, 1994; Gibson and Maunsell, 1997; Haenny et al., 1988; Maunsell et al., 1991; see also Guo and Guo, 2005 for an example in drosophila). During this task a pair of stimuli is sequentially presented, and the monkey releases a lever only if the second (test) stimulus is a match to the first (sample) stimulus. Recordings were made from neurons in areas V4 and IT – i.e. regions traditionally considered visual in their function and particularly sensitive to object features (No recordings were made in auditory cortices; something that awaits future investigation.). In these studies, selective delay-period activity was observed in response to visual-visual, somatosensory-visual, and auditory-visual paired associates. That is, this delay-period activity was observed not only in response to visual stimuli, but also non-visual stimuli and furthermore was selective for specific associations among the learned set. In these studies, like the human imaging work mentioned above, extensive studying of the paired associations was performed prior to testing. Unlike the abovementioned studies in humans, the studies in monkeys all entailed abstract forms and sounds.

One implication of these collective data is that prior multisensory experiences can influence and be part of memory functions such that when an association is formed between sensory modalities for a given object, presentation of the stimulus in just one sensory modality can alter the activity in regions typically implicated in the processing of the other, non-stimulated sensory modality. That is, responses to an incoming stimulus may vary, either in terms of their pattern within a region or overall activated network, according to whether it is part of a multisensory or unisensory memory.

2. Scope of this Review

Our research in this domain set out to address several hitherto unresolved issues (Murray et al., 2004, 2005; Lehmann and Murray, 2005; Murray and Sperdin, 2009; Thelen et al., 2012). First, we wanted to determine if multisensory experiences influence subsequent behaviour with unisensory stimuli. We would consider such an indicator of their behavioural relevance, particularly if such can

be observed with task-irrelevant multisensory experiences. Second, because prior studies all involved extensive training with or exposure to the multisensory stimuli, the requisite conditions for eliciting such effects were unknown. For example, it was unclear how much multisensory experience is necessary to observe a behaviourally-relevant effect and/or effect on brain responses. Our tactic was to use single-trial and task-irrelevant multisensory experiences during a continuous recognition task requiring old/new discrimination within one sensory modality. One collateral benefit of having participants perform the task (and presumably attend to only one sensory modality), in our opinion, was to minimize any contribution of mental imagery that may have confounded prior works requiring overt discrimination of stimuli according to whether or not they been learned in a multisensory context. That is, accurate performance of the continuous recognition task does not require participants to remember which stimuli were experienced in a multisensory context. Third, we sought to determine whether or not multisensory effects on subsequent unisensory processing were linked to differential processing of the multisensory stimuli (i.e. to differential encoding) or instead were limited to effects during the subsequent unisensory processing. Finally, we used electrical neuroimaging analyses of event-related potentials (e.g. Michel et al., 2004; Murray et al., 2008; Michel and Murray, 2012; Tzovara et al., 2012a,b) to provide adequate spatio-temporal resolution to garner information concerning when and where such effects first occur (both in terms of time post-stimulus and in terms of levels of processing).

- Figure 1 about here -

3. Paradigm at the focus of this review

To allow for investigating single-trial effects, we adopted a continuous recognition task that required participants to indicate whether or not a given stimulus was novel or had already been presented during the current block of trials. For example, the paradigm shown in Figure 1 would have participants indicating if a given image was being seen for the initial or repeated time during a block

of trials. Participants would have been instructed to perform as quickly and accurately as possible and to ignore any sounds they heard. The distribution of old/new presentations as well as unisensory and multisensory stimuli were roughly equated across quartiles of trials within a block. Over the course of our studies, we have applied several variants of this paradigm. In general, trials could be sub-divided at one level between initial and repeated presentations. Initial presentations could be further differentiated between those that were unisensory and those that were multisensory. Likewise, we were able to further vary the semantic congruence and general meaningfulness of the multisensory pairing during initial stimulus presentations. In the majority of our research, repeated presentations were only unisensory and thus could be differentiated according to the manner in which they had been initially presented (i.e. those that were always unisensory and those that had been presented in a multisensory manner and were presented subsequently in a unisensory fashion). This of course introduced a potential confound wherein multisensory stimuli always indicated initial stimulus presentations. The results provide one level of argument against this confound (discussed below). Our most recent variations of this paradigm more fully addressed this possibility by including multisensory repetitions of stimuli so that whether or not a stimulus was unisensory or multisensory was truly uninformative for the task (Thelen et al., submitted).

- Table 1 about here -

4. Psychophysical Findings

Unisensory stimuli are discriminated according to past multisensory experiences

The psychophysical findings from our collective studies are summarized in Figure 2 and Table 1. Generally-speaking, effects were limited to modulations in discrimination accuracy with no reliable effects observed on reaction times, except a general slowing for multisensory conditions versus unisensory conditions during initial object presentation (i.e. when the multisensory association was presumably being formed).

In a first study combining psychophysics and event-related potential recordings (Murray et al., 2004), multisensory conditions during initial stimulus presentations entailed semantically congruent object pairings. Stimulus repetitions were only visual, producing two conditions: repetitions that had previously been unisensory (V-) and repetitions that had previously been multisensory and semantically congruent (V+c). Participants were significantly more accurate on the V+c than V- condition (Table 1), indicating that prior single-trial multisensory experiences influence current unisensory visual processing. This was replicated and extended in a combined psychophysics and functional magnetic resonance imaging (fMRI) paradigm (Murray et al., 2005), where performance during scanning was again enhanced for the V+c versus V- condition (Table 1), despite the noise of the MRI scanner environment. It is likewise worth noting that this effect withstood the constraints of acquisitions within the scanner environment including increased background noise and the extension of the temporal lag between initial and repeated image presentations to approximately 50 seconds (while keeping the absolute number of intervening trials approximately equivalent to that in Murray et al., 2004).

Subsequent research examined the requisite conditions for observing multisensory effects on later unisensory object discrimination. On the one hand, Lehmann and Murray (2005) examined the role of semantic congruence by dividing initial image presentations into three groups: those appearing only visually (50% of initial trials; V), those appearing as a semantically congruent auditory-visual pair (25% of initial trials; AVc), and those appearing as a semantically incongruent auditory-visual pair (25% of initial trials; AVi) (their Experiment 2). This manipulation led to a significant modulation in memory performance with image repetitions (main effect $F_{(2,9)}=23.95$; $p<0.001$; $\eta_p^2=0.842$). More specifically, performance was enhanced for those images that had been paired with a semantically congruent environmental sound (V+c) relative to either those images only appearing visually (i.e. the V- condition) or images that had been paired with a semantically incongruent environmental sound (V+i) (Table 1).

On the other hand, Lehmann and Murray (2005) and more recently Thelen et al. (2012) examined the importance of using meaningful sounds in order for multisensory memories to impact subsequent visual processing. They paired half of the initial image presentations with meaningless sounds. The main difference between these studies is that while the same meaningless pure tone was used by Lehmann and Murray (2005), distinct meaningless sounds were used by Thelen et al. (2012). In both studies, this manipulation led to significant performance impairment for images that had been paired with meaningless sounds (i.e. the V+m condition) relative to images presented visually on both initial and repeated presentations (i.e. the V- condition) (Table 1). However, it is important to note that this impairment nonetheless provides an indication of differential processing of current visual information according to past multisensory vs. unisensory experiences; a point to which we return when discussing the neuroimaging findings.

Most recently, Thelen et al. (submitted) adopted a more controlled paradigm that focused on two aspects. First, multisensory contexts occurred with equal probability on initial and repeated stimulus presentations. Second, pairings with semantically congruent, incongruent and meaningless sounds were all intermixed within the same block of trials completed by the same group of participants. Analysis of the accuracy data with repeated visual stimuli revealed a main effect of condition ($F_{(3,23)}$=7.990; p=0.001; η_p^2=0.510) (Figure 2B). Post-hoc analyses revealed a pattern highly consistent with our collective prior findings (Table 1). In comparison to images only encountered in a unisensory visual context, accuracy was improved for images that had been previously presented with a semantically congruent sound and was impaired for images that had been previously presented with a semantically incongruent or meaningless sound. Once again, there was no evidence of significant effects on reaction times to repeated visual stimuli.

In addition to these changes, Thelen et al. (submitted) examined the extent to which auditory object discrimination is impacted by such task-irrelevant and single-trial multisensory experiences. Analysis of the accuracy data with repeated auditory stimuli revealed a main effect of condition ($F_{(3,23)}$=21.685; p<0.001; η_p^2=0.739) (Figure 2B). Post-hoc analyses revealed a pattern highly

consistent with our findings with images (Table 1). In comparison to sounds only encountered in a unisensory auditory context, accuracy was improved for sounds that had been previously presented with a semantically congruent image and was impaired for sounds that had been previously presented with a semantically incongruent image. There was no reliable effect of prior pairings with meaningless images. There was no evidence of significant effects on reaction times to repeated auditory stimuli.

Importance of initial multisensory encounters

It can reasonably be asked whether the above effects of multisensory memories on current unisensory recognition require multisensory experiences per se or can also be observed conversely (i.e. when initially unisensory stimuli are subsequently presented in either a unisensory or multisensory manner). This could be (partially) addressed in the modified paradigm of Thelen et al. (submitted), which entailed a 2x3 within-subject design with factors of initial presentation type (unisensory vs. multisensory) and repeated multisensory condition (semantically congruent, incongruent, and meaningless). Accuracy data revealed a main effect of initial presentation type, with generally higher accuracy in discriminating repeated images (all of which are presented as multisensory pairs) when the initial presentation had been multisensory rather than unisensory (95.5% vs. 93.2%; $F_{(1,25)}$=7.801; p=0.010; η_p^2=0.238). There was also a significant interaction between initial presentation type and repeated multisensory condition ($F_{(2,24)}$=5.256; p=0.013; η_p^2=0.305). The main effect of repeated multisensory condition was not significant ($F_{(2,24)}$=0.573; p=0.572; η_p^2=0.046). Given this significant interaction, separate 1-way ANOVAs were conducted for each initial presentation type. Accuracy did not significantly vary across repeated multisensory conditions if the initial presentation had been unisensory ($F_{(2,24)}$=0.244; p=0.786; η_p^2=0.020). By contrast, accuracy significantly differed across repeated multisensory conditions if the initial presentation had been multisensory ($F_{(2,24)}$=3.642; p=0.042; η_p^2=0.233). Performance was more accurate for semantically congruent than incongruent pairs (97.1% vs. 94.1%; $t_{(25)}$=2.658; p=0.0135); no other paired contrasts

reach the 0.05 significance criterion. These results provide an additional demonstration of the efficacy of multisensory memories on later image recognition.

Role of the number of intervening trials

We likewise examined whether the number of intervening trials influenced the efficacy of single-trial multisensory memories on later unisensory discrimination. This was done using the data from Experiment 1 of Thelen et al. (submitted). In this experiment there was a range of 5-13 intervening trials. This range was distributed across conditions such that ~50% of intervals were from 5 to 10 items (short) and ~50% of intervals were from 11 to 13 items (long). Accuracy data from the V-, V+c, V+i, and V+m conditions were submitted to a 4x2 repeated measures ANOVA with condition and interval bin (short/long) as within-subject factors. This analysis revealed a main effect of condition ($F_{(3,23)}$=7.258; p=0.001; η_p^2=0.486; see Figure 2B and Table 1) and a condition x interval bin interaction ($F_{(3,23)}$=3.071; p=0.048; η_p^2=0.286). There was no reliable main effect of bin ($F_{(1,25)}$=2.156; p=0.155; η_p^2=0.079). A set of paired t-tests comparing short vs. long intervals for each condition was performed to better isolate the bases of this interaction. Interval length only affected performance on the V+i condition (93±2% vs. 87±2%; $t_{(25)}$=2.858; p<0.009). No other condition was significantly affected (p-values >0.45). Given this pattern of results, the range of intervals we have used between initial and repeated presentations seems to have limited, if any, effect on the overall efficacy of multisensory memories on current unisensory processing.

Effects on memory performance are dissociable from encoding

Another consistency across our studies is that the patterns of performance on initial and repeated presentations were dissociable (Figure 2A). Specifically, while accuracy is affected on repeated presentations without evidence for effects on reaction time; performance on initial presentations is significantly slowed on multisensory vs. unisensory trials, irrespective of whether the auditory-visual pairings were semantically congruent, incongruent, or involved meaningless stimuli.

This was the case despite performance accuracy being equivalent (and near ceiling) across all initial presentations. Consequently, the above effects on memory performance cannot be readily explained as a direct transfer of an effect occurring during initial image presentation and multisensory encoding/interactions. However, the possibility that equivalent performance measures are nonetheless masking differential brain processes cannot be unequivocally excluded. This is something that additional brain imaging studies will need to address. Preliminary results would suggest that differences in multisensory, but not unisensory, processing during initial image presentations are predictive of whether or not subsequent memory performance will be facilitated (Thelen and Murray, 2013).

The role of attention, alerting, and novelty

We would propose that these behavioural effects follow from distinct neural representations of multisensory and unisensory experiences that are formed by single-trial exposures and later accessible during subsequent unisensory processing. Nonetheless, it is worthwhile to also consider some alternative accounts. One possibility is that these effects are the consequence of selective attention to the auditory channel and/or novel contexts (e.g. Tsivilis et al., 2001; Ranganath and Rainer, 2003). Such accounts would have predicted faster and/or more accurate performance on initial multisensory presentations, particularly because the mere presence of non-visual information would have been a sufficient cue to indicate a novel image presentation (see e.g. Chen and Yeh, 2008). That is, on the basis of selectively attending to audition, subjects would have been able to more accurately and rapidly indicate an image's initial presentation (for multisensory versus unisensory trials). Such a pattern was not observed in any of our experiments. A similar argument applies to an explanation in terms of general alerting, wherein multisensory events would have been predicted to produce the fastest behaviour. Rather, the pattern of reaction times on initial stimulus presentations fits well with results suggesting that events in an unexpected modality during a discrimination task can lead to slowed reaction times (Spence et al., 2001). However, this variety of

selective attention still would not account for the performance pattern observed with repeated image presentations, particularly those where the semantic congruence was varied (see Figure 2 and Table 1). In addition, effects of general arousal and fatigue cannot readily account for our results, because the experimental design included a nearly homogenous distribution of the different stimulus conditions throughout blocks of trials. Thus, even if subjects were more engaged in the task during the beginning of a block of trials, this would have applied equally to all stimulus conditions. While our efforts to date have been to minimize/exclude effects of attention, it would be informative to explicitly manipulate participants' attention to either the task-relevant or task-irrelevant sensory modality (e.g. by varying stimulus salience or signal-to-noise in either/both sensory modalities or by varying the task requirements).

- Figure 3 about here -

5. Brain imaging findings

Responses to visual stimuli differ during early processing stages according to past multisensory experiences

In addition to the above behavioural effects, brain responses significantly differed between image repetitions that had been initially presented in an auditory-visual multisensory context and those that had exclusively been presented visually (Murray et al., 2004, 2005; Thelen et al., 2012). In an event-related potential (ERP) study brain responses to V+c and V- conditions first differed over the 60-136ms post-stimulus period (Figure 3A) (Murray et al., 2004). Electrical neuroimaging analyses determined that this difference was the consequence of changes in the electric field topography at the scalp, rather than its amplitude (i.e. strength). That is, responses to the V+c and V- conditions differed in terms of the configuration of the generators active over this time period, such that different sets of brain regions were active at 60-136ms post-stimulus onset depending on whether or not the incoming visual stimulus had been initially encountered in a multisensory context. Source

estimations performed throughout the entire gray-matter volume and statistical analyses thereof indicated that distinct subsets of lateral occipital cortices mediated this early effect (Figure 3B). An event-related fMRI study at 1.5T with whole-brain acquisition (Murray et al., 2005) both confirmed the localization provided by the ERP source estimations and also addressed discrepancies between Murray et al. (2004) and prior hemodynamic imaging studies (most notably those of Nyberg et al., 2000 and Wheeler et al., 2000). As already detailed above, we were able to replicate our behavioural findings despite the modifications to the paradigm necessitated by fMRI constraints (i.e. the additional time between trials and the additional acoustic noise from the scanner gradients). Additionally, we replicated the observation of response modulations within lateral occipital cortices between V+c and V- conditions.[1] In both studies responses were stronger to the V+c than V- condition within the lateral occipital cortex.

Most recently, Thelen et al. (2012) used electrical neuroimaging analyses of ERPs to show that brain responses to repeated visual objects differed according to whether the initial encounter was exclusively visual or had included a meaningless sound. Similar to the findings of Murray et al. (2004), the difference in brain responses to V+m versus V- began at 100ms post-stimulus onset, were due to topographic differences in the ERP, and were localized to both a small cluster within the right lateral occipital cortex as well as a larger cluster within the right posterior superior temporal gyrus (Figure 3). Effects within the lateral occipital cortex followed from stronger responses to the V- than V+m condition; i.e. to images that had been exclusively presented in a visual context vs. those initially presented with a meaningless sound. Conversely, effects within the superior temporal gyrus followed from stronger responses to the V+m than V- condition. Later effects (270-310ms post-stimulus onset) were again the consequence of topographic ERP differences and were now localized to the right middle temporal gyrus. Activity was stronger for the V+m than V- condition; i.e. for images that had been paired with a meaningless sound. Stronger behavioural decrements were correlated with stronger differential activity within the middle temporal gyrus (detailed in Thelen et al., 2012).

Both of the above ERP studies converge on a common time window at approximately 100ms post-stimulus onset during which differential brain activity is observed as a function of prior task-irrelevant multisensory contexts. The specific network involved seems to depend on whether or not the prior multisensory experience facilitates or impairs memory performance. At least in the case of image repetitions, lateral occipital cortices respond more strongly to the condition leading to more accurate memory performance and regions within temporal cortices respond more strongly if past multisensory experiences impair memory performance; this latter finding is consistent with hemodynamic imaging results of Nyberg et al (2000) and Wheeler et al. (2000) (see also Tanabe et al., 2005).

6. Implications

Our principal finding across these studies is that past multisensory experiences influence both the ability to accurately discriminate image repetitions during a continuous recognition task as well as brain responses to image repetitions – thereby extending the effects of multisensory interactions across a substantially longer timescale than previously considered. This discrimination was according to past multisensory versus unisensory experiences, during the task itself, and was influenced by both the simple co-occurrence of an unrelated, meaningless stimulus of another sensory modality as well as semantic features (i.e. the co-occurrence of meaningful object stimuli). Accuracy in indicating visual object repetitions (1) was significantly impaired for those images that had been presented with a meaningless sound, (2) was not reliably affected for those images that had initially been presented with a semantically incongruent sound, and (3) selectively improved for images initially presented with a semantically congruent sound. Such performance changes were relative to repetition discrimination accuracy with those images initially presented only visually. This pattern generalized to when participants discriminated initial versus repeated auditory object presentations. These effects provide some indications concerning the necessary conditions for multisensory perceptual/memory traces to be established and later accessed upon the repeated presentation of

unisensory visual or auditory stimuli. The collective results reveal opposing effects of meaningful and meaningless semantic contexts from auditory-visual multisensory events.

Our findings challenge the proposal that single-trial multisensory interactions impact subsequent unisensory retrieval only in specific semantically congruent situations or when information across the senses had been concordant. Some prior works placed an emphasis on the role of object familiarity (van der Linden et al., 2010) or ethological relevance (von Kriegstein and Giraud, 2006). Likewise, these results challenge an interpretation wherein meaningless sounds/images result in a noisy representation of objects; something we had initially proposed in Lehmann and Murray (2005). Instead, the data speak in favour of distinct (though potentially short-lived) representations being established following single-trial multisensory events that in turn can impact subsequent behaviour to and processing of unisensory components of those multisensory events. Such a pattern of results likewise suggests that a certain degree of learning of new associations (or of specific associations of already-learned objects) is being formed during the experiment to impact later unisensory processing. Had this not been the case, then pairing meaningful objects with either meaningless stimuli or semantically incongruent stimuli would not have led to behavior different from that observed with stimuli only encountered in a unisensory context.

Linked to the establishment of such representations is the question of the extent to which these phenomena can be interpreted in the broader framework of redintegration (Hamilton, 1859). Redintegration refers to the capacity of a portion of a consolidated memory to re-activate the entire extended original network. Do unisensory stimuli presented subsequently to (single-trial) multisensory experiences result in redintegration? Incorporating our findings into this framework would instead suggest that redintegration processes might also manifest without explicit consolidation of auditory-visual associations and first within regions principally involved in multisensory interactions rather than with memory consolidation. The design of the continuous recognition task used in our work did not permit extensive studying of the multisensory associations.

There were only single-trial exposures, and the initial and repeated presentations were pseudo-randomly intermixed. More generally, the observed performance facilitation (and impairment) does not appear to be contingent upon extensive or explicit encoding.

Important aspects that warrant continued research concern the duration over which such effects persist as well as the *a priori* prediction of whether or not multisensory stimuli will indeed benefit later memory performance. Both of these will be particularly important for any application of this paradigm in a clinical or training setting. Our paradigms have thus far separated initial and repeated presentations of the same object by ~10 intervening items, which corresponded to delays of ~20-60 seconds. The real-world example described in the introduction of this review would, however, suggest that effects may persist over longer delays. With regard to predicting the performance effects of prior multisensory experiences, the data reviewed here consistently demonstrate relative memory improvements for stimuli that had been initially encountered in a semantically congruent multisensory context and memory impairments for stimuli that had been initially encountered in a meaningless multisensory context. While this consistency is observed when using a continuous recognition task, such has not been the case with tasks entailing separate study/test sessions and by extension explicit discrimination of the context in which a specific item had been initially encountered (Nyberg et al., 2000; Wheeler et al., 2000; von Kriegstein and Giraud, 2006; Butler and James, 2011).

Another domain warranting investigation in humans concerns the dependence of the effects reported here on object processing (at least in the task-relevant modality). While abstract designs/sounds have not yet been applied in the context of a continuous recognition task like that shown in Figure 1, such stimuli have been used in single-unit recordings within areas V4/IT during a delayed match-to-sample task (e.g. Colombo and Gross, 1994; Gibson and Maunsell, 1997; Haenny et al., 1988; Maunsell et al., 1991). The fact that selective delay activity was observed in these studies for not only visual-visual pairings, but also visual-auditory and auditory-visual pairings would suggest that the object-ness of the stimuli is not a determining factor. It should be noted, however, that the

stimuli and their associations were extensively studied (and presumably learned) by the monkeys. It is therefore not clear if such effects would occur when the associations are less familiar and based on single-trial experiences. A similar critique applies to a recent positron emission tomography (PET) study in humans (Zangenehpour and Zatorre, 2010). In this study, one group of participants was exposed (~45min) to spatially and temporally co-occurring noise bursts and LED flashes prior to the PET acquisition on the following day (group E). The other (naive) group of participants did not have this exposure (group N). The PET data indicated that auditory noise bursts resulted in increased cerebral blood flow in primary visual cortex of participants from group E but not group N. This was paralleled by increased functional connectivity between primary auditory and visual cortices. Zangenehpour and Zatorre (2010) argue this as evidence for auditory-driven activity of visual cortex as a consequence of implicitly learned associations between meaningless auditory-visual stimuli; the effects of which persisted at least 24 hours. Several issues will warrant further investigation. For example, it will be important to determine the dependence of this effect on the length of the exposure session and the contribution (if any) of mental imagery. It will likewise be important to ascertain the latency at which the auditory stimulus initiates responses within primary visual cortices. While the functional connectivity data of Zangenehpour and Zatorre (2010) argue for a direct route and presumably a short lag between auditory and visual cortical responses, this remains to be empirically demonstrated. Such points notwithstanding, the collective results from monkeys and humans would suggest that meaningless stimuli are effective in engendering multisensory representations that impact subsequent unisensory processing.

Along these lines, we should mention an additional model of multisensory object processing that postulates that information from the different senses first converges in perirhinal cortices (reviewed in E.A. Murray and Bussey, 1999). While it is undeniable that there is convergence of multisensory information in perirhinal cortices (cf. Box 3 in Murray and Bussey, 1999) and that lesions to these regions temporarily impair performance with learned multisensory associations (e.g Murray and Gaffan, 1994; see also Taylor et al., 2006, 2009), it is far less clear if this is the *first* locus

of multisensory integration and/or the most pertinent for the effects reviewed above. Indeed, evidence would suggest there to be multiple anatomic pathways, including but not limited to a direct pathway between primary sensory cortices (reviewed in Ghazanfar and Schroeder, 2006; Cappe et al., 2009). This is clearly a domain where additional research is required to disambiguate the role(s) of specific brain circuits in varieties of multisensory integration and memory performance.

7. Conclusion

In conclusion, the findings reviewed here highlight the functional efficacy of multisensory memories on performance and brain activity not only when the multisensory associations are explicitly learned, but also when such associations are formed incidentally after single-trial exposure. The growing interest in multisensory learning (e.g. Naumer et al., 2009; Shams and Seitz, 2008) and long-term effects of multisensory interactions more generally (e.g. Meylan and Murray, 2007; Naue et al., 2011; Shams et al., 2011; Wozny and Shams, 2011) is not only opening new lines of basic research, but also strategies for education and clinical rehabilitation (e.g. Johansson, 2012).

Acknowledgements

Financial support was provided by the Swiss National Science Foundation (grant no. 310030B-133136 to M.M.M).

References

Amedi, A., von Kriegstein, K., van Atteveldt, N.M., Beauchamp, M.S., and Naumer, M.J. (2005). Functional imaging of human crossmodal identification and object reocognition. *Experimental Brain Research* **166**, 559-71.

Beer, A.L., Plank, T., Meyer, G., and Greenlee, M.W. (2013). Combined diffusion-weighted and functional magnetic resonance imaging reveals a temporal-occipital network involved in auditory-visual object processing. *Frontiers in Integrative Neuroscience* **7**, 5.

Butler, A.J., and James, K.H. (2011). Cross-modal versus within-modal recall: Differences in behavioral and brain responses. *Behavioural Brain Research* **224**, 387-96.

Cappe, C., Rouiller, E.M., and Barone, P. (2009). Multisensory anatomical pathways. *Hearing Research* **258**, 28-36.

Cappe, C., Thut, G., Romei, V., and Murray, M.M. (2010). Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *Journal of Neuroscience* **30**, 12572-80

Cappe, C., Thelen, A., Romei, V., Thut, G., and Murray, M.M. (2012). Looming signals reveal synergistic principles of multisensory interactions. *Journal of Neuroscience* **32**, 1171-82.

Chen, Y.C., and Yeh, S.L. (2008). Visual events modulated by sound in repetition blindness. *Psychonomic Bulletin & Review* **15**, 404-8.

Colombo, M., and Gross, C.G. (1994). Responses of inferior temporal cortex and hippocampal neurons during delayed matching to sample in monkeys (Macaca fascicularis). *Behavioral Neuroscience* **108**, 443-55.

Ghazanfaar, A.A., and Schroeder, C.E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences* **10**, 278-85.

Gibson, J.R., and Maunsell, J.H.R. (1997). Sensory modality specificity of neural activity related to memory in visual cortex. *Journal of Neurophysiology* **78**, 1263-75.

Gottfried, J.A., Smith, A.P.R., Rugg, M.D., and Dolan, R.J. (2004). Remembrance of odors past: human olfactory cortex in cross-modal recognition memory. *Neuron* **42**, 687-95.

Goulet, S., and Murray, E.A. (2001). Neural substrates of crossmodal association memory in monkeys: the amygdala versus the anterior rhinal cortex. *Behavioral Neuroscience* **115**, 271-84.

Guo, J., and Guo, A. (2005). Crossmodal interactions between olfactory and visual learning in Drosophila. *Science* **309**, 307-10.

Haenny, P.E., Maunsell, J.H.R., and Schiler, P.H. (1988). State dependent activity in monkey visual cortex: II. Retinal and extraretinal factors in V4. *Experimental Brain Research* **69**, 245-59.

Hamilton, W. (1859). *Lectures on Metaphysics and Logic.* Gould & Lincoln, Boston, MA, USA.

James, T.W., Humphrey, G.K., Gati, J.S., Servos, P., Menon, R.S., and Goodale, M.A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* **40**, 1706–14.

Johansson, B.B. (2012). Multisensory stimulation in stroke rehabilitation. *Frontiers in Human Neuroscience* **6**, 60.

Lehmann, S., and Murray, M.M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Research Cognitive Brain Research* **24**, 326-34.

Maunsell, J.H.R., Sclar, G., Nealey, T.A., and Depriest, D.D. (1991). Extraretinal representations in area V4 in the macaque monkey. *Visual Neuroscience* **7**, 561-73.

Meylan, R.V., and Murray, M.M. (2007). Auditory-visual multisensory interactions attenuate subsequent visual responses in humans. *Neuroimage* **35**, 244-54.

Michel, C.M., and Murray, M.M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage* **61**, 371-85.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology* **115**, 2195-222.

Murray, E.A., and Bussey, T.J. (1999). Perceptual–mnemonic functions of the perirhinal cortex. *Trends in Cognitive Sciences* **3**, 142-51.

Murray, M.M., Cappe, C., Romei, V., Martuzzi, R., and Thut, G. (2012). Auditory-visual multisensory interactions in humans: a synthesis of findings from behavior, ERPs, fMRI, and TMS, in: *The*

New Handbook of Multisensory Processes, B.E. Stein (Ed.), MIT Press, Cambridge, MA, USA. pages 223-38.

Murray, M. M., and Sperdin, H. F. (2010). Single-trial multisensory learning and memory retrieval, in: *Multisensory Object Perception in the Primate Brain,* J. Kaiser and M.J. Naumer (Eds.), Springer, Heidelberg, Germany. pages 191-208.

Murray, M.M., Brunet, D., and Michel, C.M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain Topography* **20**, 249–64.

Murray, M.M., Foxe, J.J., and Wylie, G.R. (2005). The brain uses single-trial multisensory memories to discriminate without awareness. *Neuroimage* **27**, 473-8.

Murray, M.M., Michel, C.M., Grave de Peralta, R., Ortigue, S., Brunet, D., Andino, S.G., and Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* **21**, 125-35.

Naci, L., Taylor, K.I., Cusack, R., and Tyler, L.K. (2012). Are the senses enough for sense? Early high-level feedback shapes our comprehension of multisensory objects. *Frontiers in Integrative Neuroscience* **6**, 82.

Naue, N., Rach, S., Struber, D., Huster, R.J., Zaehle, T., Korner, U., Herrmann, C.S. (2011). Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *Journal of Neuroscience* **31**, 7729-36.

Naumer, M.J., Doehrmann, O., Müller, N.G., Muckli, L., Kaiser, J., and Hein, G. (2009). Cortical plasticity of audio-visual object representations. *Cerebral Cortex* **19**, 1641-53.

Nyberg, L., Habib, R., McIntosh, A.R., and Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences USA* **97**, 11120-4.

Ranganath, C., and Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience* **4**, 193-202.

Shams, L., Wozny, D. R., Kim, R., and Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology* **2**, 264.

Shams, L., and Seitz, A.R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences* **12**, 411-7.

Spence, C., Nicholls, M.E., and Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception and Psychophysics* **63**, 330-6.

Tanabe, H.C., Honda, M., and Sadato, N. (2005). Functionally segregated neural substrates for arbitrary audiovisual paired-association learning. *Journal of Neuroscience* **25**, 6409-18.

Taylor, K.I., Stamatakis, E.A., and Tyler, L.K. (2009). Crossmodal integration of object features: voxel-based correlations in brain-damaged patients. *Brain* 132, 671-83.

Taylor, K.I., Moss, H.E., Stamatakis, E.A., and Tyler, L.K. (2006). Binding crossmodal object features in perirhinal cortex. *Proceedings of the National Academy of Sciences USA* 103, 8239-44.

Thelen, A., Cappe, C., and Murray, M.M. (2012). Electrical neuroimaging of memory discrimination based on single-trial multisensory learning. *Neuroimage* **62**, 1478-88.

Thelen, A., Talsma, D., and Murray, M.M. (submitted). The efficacy of single-trial multisensory memories for visual and auditory object recognition.

Thelen, A., and Murray, M.M. (2013) Predicting individual differences in the impact of multisensory, single-trial exposure upon subsequent object recognition. Program No. 765.15483. Neuroscience 2013 Abstracts. San Diego, California: Society for Neuroscience, 2013. Online.

Tsivilis, D., Otten, L.J., and Rugg, M.D. (2001). Context effects on the neural correlates of recognition memory: an electrophysiological study. *Neuron* **31**, 497-505.

Tzovara, A., Murray, M.M., Plomp, G., Herzog, M., Michel, C.M., and De Lucia, M. (2012). Decoding stimulus-related information from single-trial EEG responses based on voltage topographies. *Pattern Recognition* **45**, 2109-22.

Tzovara, A., Murray, M.M., Michel, C.M., and De Lucia, M. (2012). A tutorial review of electrical neuroimaging from group-average to single-trial event-related potentials. *Developmental Neuropsychology* **37**, 518-544.

van der Linden, M., van Turennout, M., and Indefrey, P. (2010). Formation of category representations in superior temporal sulcus. *Journal of Cognitive Neuroscience* **22**, 1270-82.

Von Kriegstein, K., and Giraud, A.L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biology* **4**, e326.

Wheeler, M.E., Petersen, S.E., and Buckner, R.L. (2000). Memory's echo: vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences USA* **97**, 11125-9.

Wozny, D. R., and Shams, L. (2011). Computational characterization of visually induced auditory spatial adaptation. *Frontiers in Integrative Neuroscience* **5**, 75.

Footnotes

[1]While the statistical threshold used in both the ERP and fMRI study identified lateralized effects (right hemisphere in Murray et al. 2004 and left hemisphere in Murray et al. 2005), slightly more lax criteria identified bilateral clusters within the lateral occipital cortex.

Table 1. Summary of psychophysical results.

Study	Contrast		Δ%	*N*	*t*	*p*	η_p^2
Murray et al. (2004)	***V-***	***V+c***					
	86.1%	88.5%	2.4	11	$t_{(10)}$=3.18	0.010	0.50
Murray et al. (2005)	***V-***	***V+c***					
	87.8%	90.2%	2.4	8	$t_{(7)}$=2.76	0.028	0.52
Lehmann & Murray (2005) *Expt. 2*	***V-***	***V+c***					
	78.0%	83.6%	5.6	11	$t_{(10)}$=4.01	0.002	0.62
		V+i					
		77.1%	-0.9	11	$t_{(10)}$=0.41	0.650	0.02
	V+i	***V+c***					
	77.1%	83.6%	6.5	11	$t_{(10)}$=5.04	0.001	0.72
Lehmann & Murray (2005) *Expt. 1*	***V-***	***V+m***					
	79.3%	76.6%	-2.7	16	$t_{(15)}$=2.24	0.041	0.25
Thelen et al. (2012)	***V-***	***V+m***					
	87.0%	83.5%	-3.5	22	$t_{(21)}$=2.38	0.027	0.67
Thelen et al. (submitted) *Expt. 1*	***V-***	***V+c***					
	92.5%	94.9%	2.4	26	$t_{(25)}$=2.36	0.027†	0.14
		V+i					
		88.6%	-3.9	26	$t_{(25)}$=2.72	0.012†	0.19
		V+m					
		89.7%	-2.8	26	$t_{(25)}$=2.38	0.026†	0.14
	V+c	***V+i***					
	94.9%	88.6%	-6.3	26	$t_{(25)}$=4.55	0.001†	0.45
		V+m					
		89.7%	-5.2	26	$t_{(25)}$=3.19	0.002†	0.29
	V+i	***V+m***					
	88.6%	89.7%	1.1	26	$t_{(25)}$=0.59	0.280†	0.01
Thelen et al. (submitted) *Expt. 2*	***A-***	***A+c***					
	67.7%	74.0%	6.3	26	$t_{(25)}$=3.24	<0.004	0.30
		A+i					
		56.5%	-11.2	26	$t_{(25)}$=6.25	<0.001	0.61
		A+m					
		67.6%	0.00	26	$t_{(25)}$=0.17	0.949	0.00
	A+c	***AV+i***					
	74.0%	56.5%	-17.5	26	$t_{(25)}$=8.06	<0.001	0.72
		A+m					
		67.6%	-6.4	26	$t_{(25)}$=3.88	<0.001	0.38
	A+i	***A+m***					
	56.5%	67.6%	11.1	26	$t_{(25)}$=6.46	<0.001	0.63

This table lists the principal experimental conditions contrasted (contrast), the mean difference in percent correct responses for each contrast (Δ), the number of participants (*N*), as well as the t-test result, corresponding p-value, and effect size (t, p, and η_p^2, respectively). V- refers to repeated visual

stimuli that had been initially encountered in a unisensory visual context. V+c refers to repeated visual stimuli that had been initially encountered in a semantically congruent multisensory context. V+i refers to repeated visual stimuli that had been initially encountered in a semantically incongruent multisensory context. V+m refers to repeated visual stimuli that had been initially encountered in a multisensory context wherein the sound was meaningless. A homologous nomenclature applies to A-, A+c, A+i, and A+m. An † indicates the result of a 1-tailed test, which was used only in Experiment 1 of Thelen et al. (submitted) where strong *a priori* hypotheses were available.

Figure Captions

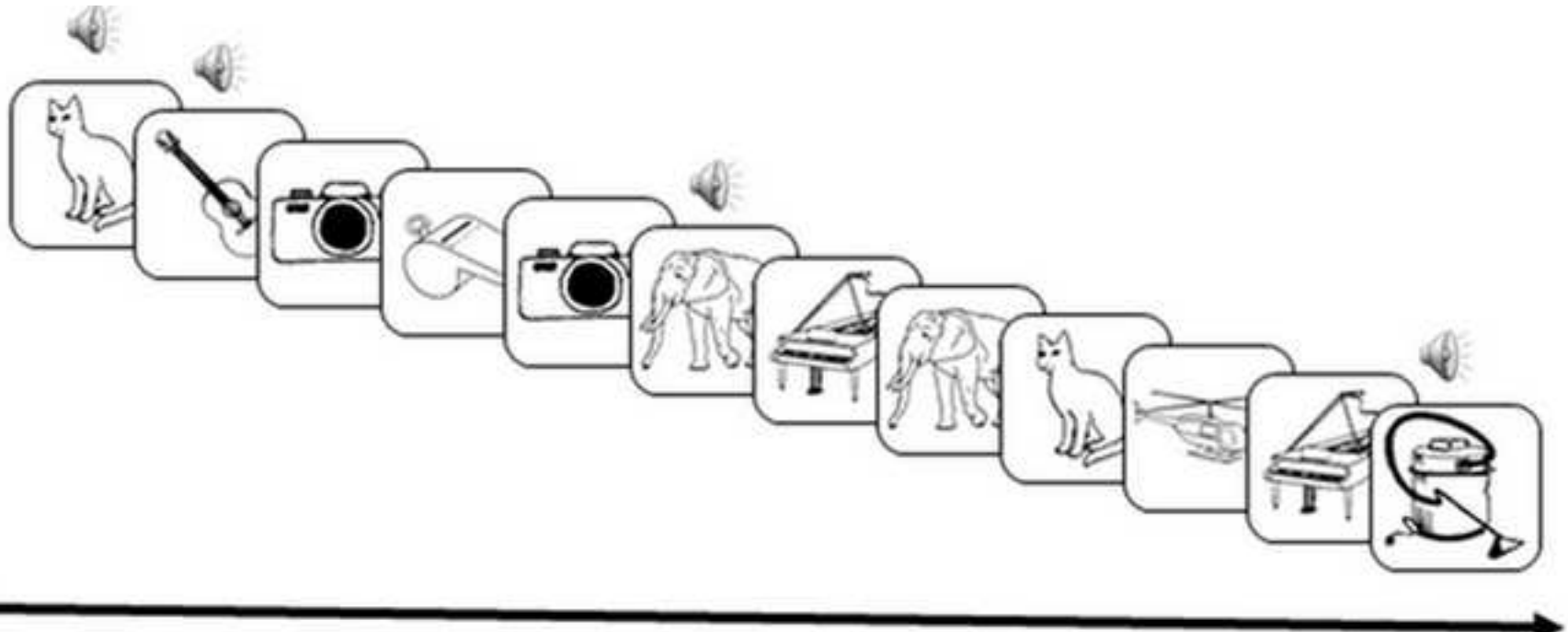
Figure 1. Illustration of the continuous recognition task used in our studies. In this paradigm participants indicate whether each image is being presented for the first or repeated time. Stimuli are presented for 500 milliseconds. Initial presentations are divided between those containing only images (V condition) and those presented with sounds (AV condition). Repeated presentations consist only of images, but can be divided between those that had been initially presented as images only (V- condition) and those that had been initially presented with sounds (V+ condition). In this way, contrasting performance and/or brain activity from the V- and V+ conditions reveals effects of past multisensory experiences on current unisensory (visual) processing.

Figure 2. Psychophysical results. Panel A: The top set of bar graphs displays the mean (s.e.m. indicated) accuracy rates on the continuous recognition task for each experimental condition. The bottom set of bar graphs displays the mean (s.e.m. indicated) reaction times. An asterisk indicates a significant difference ($p<0.05$) either for repeated presentations in the case of accuracy (see Table 1 for details) or initial presentations in the case of reaction times (details available in original publications). Panel B: The bar graphs display the mean (s.e.m. indicated) accuracy rates from Experiments 1 and 2 in Thelen et al. (submitted). In Experiment 1, 1-tailed post-hoc comparisons were warranted, while in Experiment 2 two-tailed post-hoc comparisons were used. An asterisk indicates a significant difference vs. all other conditions ($p<0.05$; see Table 1 for details). The same color across histograms refers to the same condition from different experiments.

Figure 3. Brain imaging results. Panel A displays group-averaged event-related potential waveforms from an exemplar posterior scalp site from the data of Murray et al., 2004 (left) and Thelen et al., 2012 (right). The asterisk highlights differences observed at ~100ms post-stimulus onset. The topographic maps accounting best for each condition are displayed below the waveform plots. Red indicates positive voltages, and blue negative voltages. The nasion is positioned upward and left hemiscalp on the left. Although subtle, topographic differences in each study were statistically reliable. Panel B displays the results of statistical analyses of source estimations in Murray et al.

(2004) and Thelen et al. (2012) during the earliest period of event-related potential differences as well as the results of statistical contrasts in the fMRI study of Murray et al. (2005).

Figure 1



Correct Response	new	new	new	new	old	new	new	old	old	new	old	new
Condition	AV	AV	V	V	V-	AV	V	V+	V+	V	V-	AV

time

Figure 2

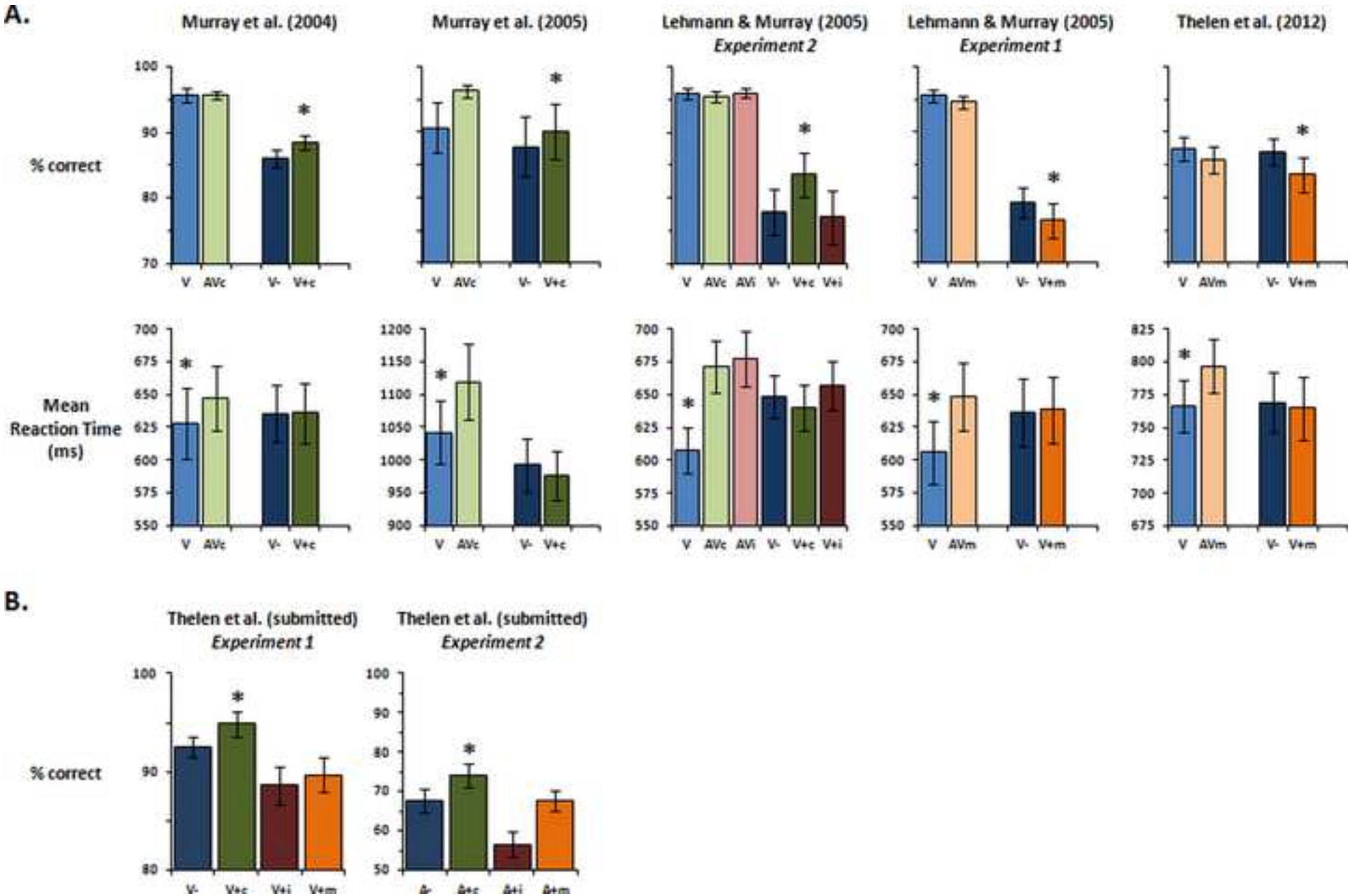
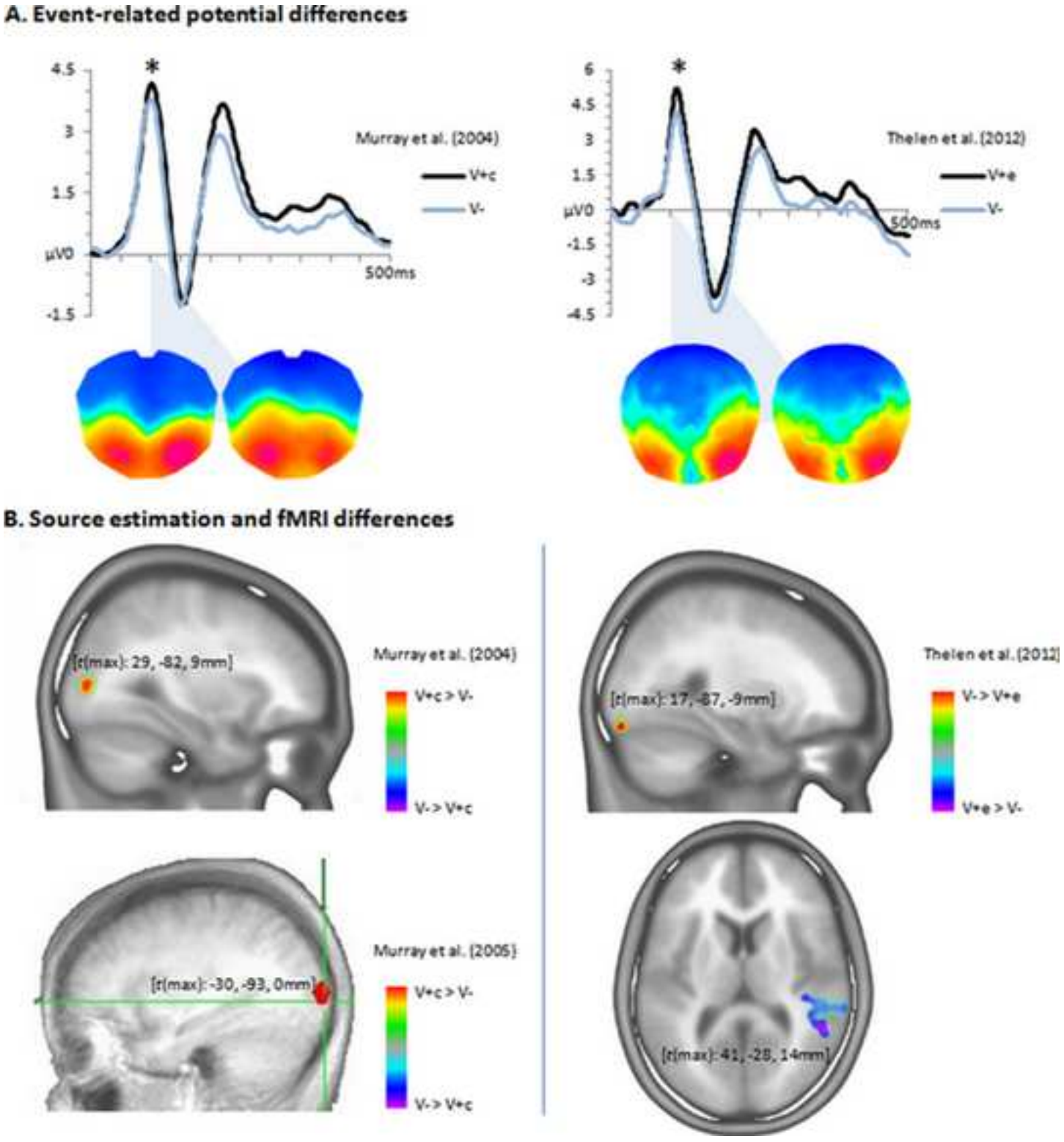

A.
Murray et al. (2004)
Murray et al. (2005)
Lehmann & Murray (2005)
Experiment 2
Lehmann & Murray (2005)
Experiment 1
Thelen et al. (2012)
% correct
Mean Reaction Time (ms)
V AVc V- V+c
V AVc AVi V- V+c V+i
V AVm V- V+m
B.
Thelen et al. (submitted)
Experiment 1
Thelen et al. (submitted)
Experiment 2
% correct

Figure 3


A. Event-related potential differences
4.5
3
1.5
µV0
-1.5
*
500ms
Murray et al. (2004)
V+c
V-
6
4.5
3
1.5
µV0
-1.5
-3
-4.5
*
500ms
Thelen et al. (2012)
V+e
V-
B. Source estimation and fMRI differences
[t(max): 29, -82, 9mm]
Murray et al. (2004)
V+c > V-
V- > V+c
[t(max): -30, -93, 0mm]
Murray et al. (2005)
V+c > V-
V- > V+c
[t(max): 17, -87, -9mm]
Thelen et al. (2012)
V- > V+e
V+e > V-
[t(max): 41, -28, 14mm]